



[3510-16-P]

DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

[Docket No.: PTO-P-2016-0039]

Standard ST.26 - Request for Comments on the Recommended Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings Using XML (eXtensible Markup Language)

AGENCY: United States Patent and Trademark Office, Commerce.

ACTION: Request for comments.

SUMMARY: Patent applications that contain disclosures of nucleotide and/or amino acid sequences must contain sequence information in a separate part of the disclosure in a specified manner. The United States Patent and Trademark Office (Office) is seeking additional comments to obtain views of the public on the continuing international effort to revise the World Intellectual Property Organization (WIPO) standard for the presentation of nucleotide and/or amino acid sequences and the consequent changes to the United States rules of practice. The revised standard will be known as WIPO Standard ST.26. An interim version of WIPO Standard ST.26 was adopted in March

2016 by the Committee on WIPO Standards (CWS), but has not been implemented pending further consideration by the CWS. Since the adoption of the interim version, efforts have been undertaken to finalize WIPO Standard ST.26 and to improve its effectiveness once implemented. One aspect of that continuing effort is a proposed guidance document annex, which will include a variety of sequence disclosure examples, to ensure understanding and uniform application of standard requirements. Comments may be offered on any aspect of this effort, and in particular, a) the comprehensiveness and clarity of WIPO Standard ST.26 and the proposed guidance document annex, and b) the proposed authoring/validation tool for creation of a sequence listing in XML.

DATES: Written comments must be received on or before [insert 60 days after date of publication in the FEDERAL REGISTER] to ensure consideration. No public hearing will be held.

ADDRESSES: Comments concerning this notice should be sent by electronic mail message over the Internet addressed to seq_listing_xml@uspto.gov. Comments may also be submitted by mail addressed to: Mail Stop Comments - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313–1450, marked to the attention of Susan C. Wolski, Office of International Patent Legal Administration, Office of the Deputy Commissioner for International Patent Cooperation. Although comments may be submitted by mail, the Office prefers to receive comments via the Internet.

The comments will be available for public inspection at the Office of the Commissioner for Patents, located in Madison East, Tenth Floor, 600 Dulany Street, Alexandria, Virginia, and will be available via the Internet (<http://www.uspto.gov>). Because comments will be made available for public inspection, information that the submitter does not desire to make public, such as an address or phone number, should not be included in the comments.

FOR FURTHER INFORMATION CONTACT: Susan C. Wolski, Office of International Patent Legal Administration, Office of the Deputy Commissioner for International Patent Cooperation, by telephone at (571) 272-3304, or by mail addressed to: Mail Stop Comments - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, marked to the attention of Susan C. Wolski.

SUPPLEMENTARY INFORMATION:

1. Background Information.

Patent applicants are currently required to submit biological sequence data in a standardized electronic format in accordance with WIPO Standard ST.25, both within the framework of the Patent Cooperation Treaty (PCT) (Annex C of the Administrative Instructions) and under most national and regional provisions. The Rules of Patent Practice in the United States (37 CFR §§ 1.821-1.825) were amended to implement WIPO Standard ST.25 in July of 1998. See Requirements for Patent Applications

Containing Nucleotide Sequences and/or Amino Acid Disclosures, 63 FR 29620 (June 1, 1998), 1121 Off. Gaz. Pat. Office 82 (June 23, 1998).

WIPO Standard ST.25, which became effective in 1998 and has not been revised since that time, requires a flat file structure of numeric identifiers using a limited set of character codes. In October 2010, the CWS established a Task Force, designating the European Patent Organization as the lead, to draft a revised standard (WIPO Standard ST.26) for the filing of nucleotide and/or amino acid sequence listings in XML format. The Office issued a first request for comments on WIPO Standard ST.26 as drafted by the Task Force (see Request for Comments on the Recommendation for the Disclosure of Sequence Listings Using XML (Proposed ST.26), 77 FR 28541 (May 15, 2012), 1379 Off. Gaz. Pat. Office 106 (June 12, 2012)), following which the draft was revised in response to comments received. In March 2016, the reconvened fourth session of the CWS adopted an interim version of WIPO Standard ST.26, which had been initially considered at the fourth session of the CWS in May 2014. The interim version of WIPO Standard ST.26 contains an editorial note requesting that implementation be postponed until the recommendation for the transition from WIPO Standard ST.25 to WIPO Standard ST.26 is agreed on by the CWS at its next session to be held in 2017. Meanwhile, WIPO Standard ST.25 should continue to be used. The adopted interim version of WIPO Standard ST.26 is composed of six documents, namely, the main body of the standard, a first annex setting forth the controlled vocabulary for use with the sequence part of the standard, a second annex setting forth the Document Type Definition (DTD) for the standard, a third annex containing a sequence listing specimen, a fourth

annex setting forth the character subset from the Unicode Basic Latin Code Table, and a fifth annex setting forth additional data exchange requirements for patent offices, and can be found here: <http://www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf>. Since the adoption of the interim version, the main body, the controlled vocabulary, and the DTD have been further revised and updated. In addition, a sixth annex has been proposed; it would contain a guidance document that aims to ensure that all applicants and Intellectual Property Offices (IPOs) understand and agree on the requirements for inclusion and representation of sequence disclosures. In all, seven rounds of discussion have been completed since March 2011, and currently, the eighth round of discussion of the documents is ongoing.

2. Request for Comments.

The Office, leading the negotiations for the United States, is seeking public comment on WIPO Standard ST.26, as revised subsequent to the adoption of the interim version. To that end, the current revisions of the main body of the standard and its five annexes, as well as the newly proposed sixth annex, are available via the Office's Web site at <http://www.uspto.gov/patent/laws-and-regulations/comments-public/2016-comments-standard-st26-presentation-nucleotide-and>.

Written comments may be offered on any aspect of WIPO Standard ST.26, its annexes, or the proposed authoring/validation tool. Comments are specifically requested on the following issues:

a) WIPO Standard ST.26 Main Body.

Since the first request for comments, the main body of WIPO Standard ST.26 has been revised, inter alia, to define a “nucleotide” to include nucleotide analogues and to provide further guidance on representation of nucleotide analogue sequences and variant sequences that have been disclosed as a single sequence with enumerated alternative variant residues at one or more positions.

The Office invites comments on whether the main body of WIPO Standard ST.26 is sufficiently comprehensive and clear, and in particular welcomes suggestions to add details or clarify the language as appropriate.

b) Guidance Document.

One goal of the development of a WIPO Standard for sequence listings is to allow patent applicants to draw up a single sequence listing in a patent application that would be acceptable for the purposes of both international and national or regional prosecution worldwide. Any new standard should represent the maximum requirements for any sequence listing submission. The purpose of the guidance document is to ensure that all applicants and IPOs understand and agree on the requirements for inclusion and representation of sequence disclosures, such that this purpose is realized.

The guidance document is composed of an introduction, examples, and a sequence listing in XML demonstrating representation of the exemplified sequences. The introduction

defines terminology used in the document and discusses the questions raised for each example, namely, whether inclusion is required for a particular disclosed sequence, if inclusion of the sequence is permitted when it is not required, and the appropriate means of representation of sequences included in a sequence listing. Examples were chosen to illustrate various paragraphs of the main body and include 22 involving nucleotide sequences and 19 involving amino acid sequences. It is envisioned that the guidance document would be updated as necessary to include further examples to keep pace with technological advances.

The Office invites comments on whether the guidance document is sufficiently comprehensive and clear, and in particular welcomes suggestions to add details or further examples as appropriate.

c) Authoring and Validation Tool.

Availability of an authoring tool in advance of the WIPO Standard ST.26 effective date is key to a successful transition from WIPO Standard ST.25. As envisioned, the authoring tool should be capable of intake of a sequence listing in WIPO Standard ST.25 format, and with additional input from applicant, create a sequence listing in WIPO Standard ST.26 format. Unfortunately, direct conversion from one standard to the other is not possible, due to numerous differences between the two standards, including inter alia, the types of required sequences, representation and annotation of the sequences, and sequence data structure.

The authoring tool should also prompt entry of all required data, prevent entry of sequences having fewer than ten specifically defined nucleotides or fewer than four specifically defined amino acids, inform as to the possibility of optional annotations, and allow use of only acceptable values or formats where applicable, thereby enhancing submission quality. A sequence listing in WIPO Standard ST.26 XML format is not as easily human-readable as its ST.25 counterpart; therefore, the tool should also provide a means for easily viewing both the in-progress and completed sequence listing.

Because the authoring tool is expected to prompt entry of all required data and to allow use of only acceptable values or formats where applicable, a certain level of validation occurs as data is entered. The tool is further expected to include a separate validation function for use by both applicants and IPOs.

WIPO Standard ST.25 provides for a single numeric identifier <223> per sequence to contain “free text” to describe sequence characteristics using non-language neutral vocabulary. Such “free text” is required to be repeated in the main part of the application description in the language thereof in a specific recommended section entitled “Sequence Listing Free Text.” Such repetition ensures that any “free text” will be translated together with the application description, precluding the need for separate translation of the sequence listing itself. In contrast, WIPO Standard ST.26 allows use of “free text” as the value for multiple different annotation qualifiers per sequence, and due to the absence of procedural requirements, repetition in the application is not required, although such a requirement under the PCT and by various IPOs is possible. In WIPO Standard ST.26,

“free text” is limited to a few short terms indispensable for understanding a characteristic of a sequence, is preferably in the English language, and as part of the sequence data part of the sequence listing, must not exceed 1000 characters composed of printable characters from the Unicode Basic Latin code table. It is expected that most inventors providing sequence information are capable of providing “free text” in the English language.

The Office invites comments on any aspect of the authoring tool, and in particular welcomes feedback on whether it is deemed necessary for the authoring tool to include a mechanism for automatic identification and extraction of any “free text” from sequence annotations to facilitate inclusion in the application description.

Dated: October 21, 2016.

Michelle K. Lee
Under Secretary of Commerce for Intellectual Property and
Director of the United States Patent and Trademark Office
[FR Doc. 2016-25968 Filed: 10/26/2016 8:45 am; Publication Date: 10/27/2016]